



Performance study of change-point detection thresholds for cumulative sum statistic in a sequential context

Nassim Sahki, Anne Gégout-Petit, Sophie Wantz-Mézières

► To cite this version:

Nassim Sahki, Anne Gégout-Petit, Sophie Wantz-Mézières. Performance study of change-point detection thresholds for cumulative sum statistic in a sequential context. *Quality and Reliability Engineering International*, 2020, 1-21, pp.21. 10.1002/qre.2723 . hal-02389331v2

HAL Id: hal-02389331

<https://inria.hal.science/hal-02389331v2>

Submitted on 11 Jul 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ARTICLE TYPE

Performance Study of Change-point Detection Thresholds for CUSUM Statistic in a Sequential Context.

Nassim Sahki* | Anne Gégout-Petit | Sophie Wantz-Mézières

¹Université de Lorraine, CNRS, Inria, IECL,
F-54000 Nancy, France.

Correspondence

*N. Sahki. Email: nassim.sahki@inria.fr /
nassim.sahki@univ-lorraine.fr /
nassimsahki06@gmail.com.

Present Address

Université de Lorraine, Institut Elie Cartan,
INRIA Nancy BIGS B.P. 70239, 54506
Vandoeuvre-lès-Nancy Cedex, France.

Abstract

Sequential detection is based on a recursive statistic and a threshold it must reach to report a change. In this paper, we consider the score-based CUSUM statistic and propose to evaluate the detection performance of some thresholds on simulated data. Three thresholds come from the literature: the Wald's constant, the empirical constant and the conditional empirical instantaneous (the latter two are built by a simulation-based procedure). Two new thresholds are built by a simulation-based procedure: the first one is instantaneous, the second is a dynamical version of the previous one. The thresholds' performance measured by an estimation of the Mean Time Between False Alarm (*MTBFA*) and the Average Detection Delay (*ADD*), are evaluated on independent and autocorrelated data for several scenarios, according to the detection objective and the real change in the data. The simulations allow us to compare the difference between the thresholds' results and to see that their performances prove to be robust when a parameter of the pre-change regime is mis-estimated or when the data independence assumption is violated. We found also that the conditional empirical threshold is the best at minimizing the detection delay while maintaining the given false alarm rate. However, on real data, we suggest to use the dynamic instantaneous threshold because it's the easiest to build for practical implementation.

KEYWORDS:

Sequential change-point detection, CUSUM Statistic, empirical threshold, simulation-based procedure

1 | INTRODUCTION

In 1931, Shewhart¹ is the first one to tackle the change-point detection problem in the general context, both offline and online. In the offline context, which is the simplest setting, the dataset is fixed and processed in one go. The goal is then the detection accuracy of change-points. Neyman & Pearson² in 1933 proposed an optimal solution that maximizes the probability of detection. In the online context, the data arrive in real time, either by point or by batch. In this case, we have to process the data sequentially in order to detect a change-point as soon as possible; detection delay is the primary concern. Sequential analysis was born with Wald's³ work in 1945, where he proposed the Sequential Probability Ratio Test (SPRT) algorithm. A test statistic is calculated at each time to decide if a change-point has occurred or not. In 1954, Page⁴ inspired by Wald's algorithm, proposed the popular CUSUM statistic ("CUmulative SUM"). The idea is to test sequentially the existence of a change-point by the recursive writing of the detection statistic as a function of the Log-Likelihood Ratio (LLR). The optimality of the CUSUM that

minimizes the average detection delay for a given bound of false alarm rate, is established by Lorden⁵ in 1971 in the asymptotic case, and by Moustakides⁶ in 1986, in the non-asymptotic case. In this sequential analysis context, another popular procedure known as Shiryaev - Roberts (Shiryaev^{7,8}; Roberts⁹) is also optimal. An asymptotic optimality was found by Pollak¹⁰ in 1985, and thereafter in 2010, a strict optimality is proposed by Polunchenko & Tartakovsky¹¹.

In this paper, we address the problem of change-point detection by designing a sequential hypothesis test. We focus on the online case where the series is supposed to be observed sequentially.

We are therefore interested in the Quickest Changepoint Detection (QCD) that minimizes the average detection delay while maintaining a given probability of false alarm. A comprehensive overview of the sequential (quickest) change-point detection problem is given by Basseville & Nikiforov¹², Poor & Hadjiladis¹³, and Tartakovsky et al.¹⁴. The sequential change-point detection is widely applied in different fields including industrial quality control¹⁵, computer network and information security^{16,17,18,19}, finance²⁰, integrity monitoring in navigation satellite systems^{14,21,22}, healthcare surveillance and epidemiology^{23,24}.

Sequential detection is based on the choice of a recursive statistic and the threshold it must reach to signal a detection. In this paper, we use the semi-parametric version of the classical CUSUM statistic given by the replacement of the LLR through a score function as suggested by Tartakovsky et al.^{16,18}. Most of the works in the literature used a constant threshold based on Wald inequality³, by controlling the false alarm rate all along the trajectory. Recently, Tsiamyrtzis & Hawkins²⁵ suggest a constant empirical threshold from simulations of the pre-change regime. Margavio et al.²⁶ in 1995, are the first to use an instantaneous detection threshold by controlling the conditional false alarm rate at each time (i.e. the probability that the current observation will trigger a false alarm, knowing that it is the first one from the beginning). It was used by Hawkins et al.²⁷ in 2003 and recently by Huang et al.²⁸ in 2016.

In our work, we propose new detection thresholds built by an empirical simulation-based method, which is similar to that of Shen et al²⁹. Our aim is to study their efficiency and to compare them with the Wald's constant, the empirical constant and the conditional empirical instantaneous thresholds on simulated data.

We recall in section 2 the context of detection: test statistic and stopping rule. In section 3, we present the different thresholds and the corresponding algorithms to build them. In section 4, we set the simulation design and the estimation procedure to evaluate the thresholds' performance. The simulation results are presented in section 5.

2 | SEQUENTIAL DETECTION

Here we give the context of the sequential detection. We consider a random phenomenon that we observe at every time. In the beginning, the behaviour is supposed to be normal. But a change that affects the behaviour of the phenomenon occurs at an unknown time. Let us describe it mathematically. Let $\{x_t\}_{t=1,\dots,n} = \{x_1, \dots, x_n\}$ a series of observations (data sequence) sequentially observed until time n not fixed. x_n is the last point added to the data set. We suppose that in the normal state, the random variables X_t 's are identically distributed (not necessary independent) according to the density distribution $f_0(\cdot)$. When a change-point occurred, this distribution changes and satisfies a post-change distribution $f_1(\cdot) \neq f_0(\cdot)$. We suppose that there exists only one change-point at time v , such that $\{X_1, \dots, X_{v-1}\}$ are identically distributed according to the pre-change distribution f_0 , and $\{X_v, X_{v+1}, \dots, X_n\}$ according to the post-change distribution f_1 .

Statistically, for all $n \geq 1$, the issue of change-point detection is to sequentially test the null hypothesis $H_{0,n}$: "no change-point occurred before time n , that is $v > n$ ", against the alternative hypothesis $H_{1,n}$: "a change-point occurred at the instant, that is $v \leq n$ " (1).

$$\begin{cases} H_{0,n} : v > n, & X_t \sim f_0(\cdot) & \forall t = 1, \dots, n. \\ H_{1,n} : \exists v \leq n, & X_t \sim f_0(\cdot) & \forall t = 1, \dots, (v-1); \\ & X_t \sim f_1(\cdot) & \forall t = v, \dots, n. \end{cases} \quad (1)$$

It is clear that there are two types of errors: the first one is to reject $H_{0,n}$ when it is true; it is a false alarm. The second one is fail to reject $H_{0,n}$, when $H_{1,n}$ is true, and it is a missed detection. Let us denote $\mathbb{P}_0[\cdot]$, $\mathbb{P}_1[\cdot]$ respectively the probability under the pre-change and post-change regime. We note by $\mu_0 = \mathbb{E}_0(X_t)$, $\sigma_0^2 = \mathbb{V}_0(X_t)$ respectively the mean and variance in the pre-change regime, and by $\mu_1 = \mathbb{E}_1(X_t)$, $\sigma_1^2 = \mathbb{V}_1(X_t)$ respectively the mean and variance in the post-change regime.

2.1 | CUSUM Recursive statistic based-score

When the two distributions f_0 and f_1 are known, it is natural to use the Log-Likelihood Ratio (LLR) to build the CUSUM statistic W_t as in Page⁴; it is defined by the following equations:

$$L_t = \log \left(\frac{f_1(x_t)}{f_0(x_t)} \right), \quad t \geq 1. \quad (2)$$

$$W_t = \max\{0, W_{t-1} + L_t\}, \quad t \geq 1, \quad W_0 = 0. \quad (3)$$

It can be noted that W_t is adapted to the online context because it is defined recursively. Notice also that in the pre-change regime, the cumulative sum of L_t up to time t is negative in expectation (because $\mathbb{E}_0(L_t) < 0$), therefore the statistic W_t is often null. In the post-change regime, from the instant of the change-point, $\mathbb{E}_1(L_t) > 0$ so that the cumulative sum of L_t increases on average and begins to become positive.

In practice, however, the two distributions f_0 and f_1 are not always known. In this case, any approach based on likelihood is useless. Tartakovsky et al. proposed to replace the LLR by a score function calculated according to observations (X_1, \dots, X_t) . They first proposed in 2006 a score function for a change-point detection on the mean in Tartakovsky et al.¹⁷. Subsequently, a generalization for detection for the mean and/or variance was given in Tartakovsky et al. in 2012¹⁸. They proposed to replace the LLR defined in (2) by a score function defined by :

$$S_t(\delta, q) = C_1 \cdot Y_t + C_2 \cdot Y_t^2 - C_3, \quad (4)$$

where $Y_t = (X_t - \mu_0)/\sigma_0$ is the centered and standardized data at time t under pre-change regime, and

$$C_1 = \delta \cdot q^2, \quad C_2 = \frac{1-q^2}{2}, \quad C_3 = \frac{\delta^2 \cdot q^2}{2} - \log(q),$$

with $\delta = (\mu_1 - \mu_0)/\sigma_0$ and $q = \sigma_0/\sigma_1$.

The parameters δ and q are set according to the detection objective which concerns the mean and/or the variance. That is, in practice, one needs to know the expected type and level of change-point to be able to set these parameters:

- δ is the mean difference between the post-change and pre-change regime normalized by the pre-change standard deviation. It can be considered as the minimum level of change on the average that we want to detect. Note that if the change-point we are looking for only concerns the variance: $\mu_1 = \mu_0$, then $\delta = 0$, and $C_1 = 0$.
- q is the ratio of variances between the pre-change and post-change regime. It is considered as the minimum level of variance change that we want to detect. If the change-point we are looking for only concerns the mean (no detection on the variance): $\sigma_1 = \sigma_0$, then $q = 1$, so $C_2 = 0$.

Following Equation (3), the based-score CUSUM statistic is defined recursively at time t as follows :

$$W_t(\delta, q) = \max\{0, W_{t-1}(\delta, q) + S_t(\delta, q)\}, \quad t \geq 1; \quad W_0(\delta, q) = 0. \quad (5)$$

The W-statistic (5) based on the score can be defined according to the knowledge of the mean and variance in the pre-change regime (μ_0, σ_0^2) and of the detection objective $(\delta$ and $q)$.

Remark: One can easily show that, like the LLR L_t , the statistic S_t satisfies $\mathbb{E}_0(S_t) < 0$ and $\mathbb{E}_1(S_t) > 0$. Note also that if the pre-change and post-change distributions of the data (f_0 and f_1) are Gaussian with known parameters μ_1, σ_1^2 and μ_0, σ_0^2 , the two statistics S_t and L_t are equal.

2.2 | Stopping rule

Under post-change regime, knowing that the W-statistic tends to grow gradually, it is natural to reject $H_{0,t}$ when the statistic W_t exceeds a threshold. This threshold has to be chosen according to an objective of false alarm rate that we denote by α . The stopping rule is to trigger an alarm in order to signal that a change-point has occurred before time t when $W_t(\delta, q)$ exceeds the detection threshold $h(\alpha)$ set in advance⁴. An alarm occurs at a time T defined as follows:

$$T_{h(\alpha)} = \min\{t \geq 1 : W_t(\delta, q) \geq h(\alpha)\}. \quad (6)$$

We note that the choice of a statistic and a stopping rule define a detection procedure that can be evaluated under the pre- and post-change regimes. Note that T is a stopping time according to the filtration generated by the $(W_t)_{t \geq 0}$ and it is an estimator of the real change point v . These different cases may arise:

- $T \geq v$, the change-point v is detected with a delay $(T - v)$;
- $T < v$, the change-point v has not yet occurred at time T . It is said that the procedure triggered a false alarm;
- $v < +\infty$, $T = +\infty$, the change-point v is not detected. This is a missed detection.

Page⁴ and Lorden⁵ propose the Mean Time Between False Alarm (MTBFA) parameter to evaluate the average number of observations before triggering a false alarm. It is given by the expectation under pre-change regime of the first alarm time:

$$MTBFA = \mathbb{E}_0 [T]. \quad (7)$$

The higher the *MTBFA*, the lower the real risk of false alarm. The instantaneous false alarm rate (or risk) is the inverse of the *MTBFA*:

$$\alpha = \frac{1}{\mathbb{E}_0 [T]}. \quad (8)$$

The Average Detection Delay (*ADD*) is evaluated under post-change regime to quantify the speed of a detection (Pollak¹⁰). It is given by the expectation under post-change regime of the stopping time T :

$$ADD = \mathbb{E}_1 [T]. \quad (9)$$

The *ARL* (Average Run Length) function¹² is used in sequential framework to unify both criteria *MTBFA* and *ADD*. It defines the *MTBFA* under pre-change regime, and the *ADD* under post-change regime.

3 | DETECTION THRESHOLDS

In this section, we first present three different thresholds from the literature: the constant threshold based on the Wald inequality³, an empirical constant threshold identical to Tsiamyrtzis'²⁵ and an empirical instantaneous threshold proposed by Margavio²⁶. Thereafter, we present the proposed two new empirical thresholds.

3.1 | Wald Constant Threshold

In the sequential detection context, the conventional threshold used in the literature is constant over time. That means that we would like to control the false alarm rate along all the trajectory $(X_1, \dots, X_n)_{n \geq 1}$. The constant threshold $h(\alpha)$ is determined such that the risk to detect an alarm at time t is controlled by α :

$$\mathbb{P}_0 [W_t(\delta, q) \geq h(\alpha)] \leq \alpha, \quad t \geq 1. \quad (10)$$

The classical method^{5,12,16,18,22} used to set a constant threshold is based on the following Wald's inequality³: $\alpha \leq e^{-h}$. Therefore, the Wald Constant Threshold (WC-Threshold) $h^W(\alpha)$ is given after fixing the tolerated α , by respecting:

$$h^W(\alpha) = -\ln(\alpha). \quad (11)$$

3.2 | Empirical Constant Threshold

The Empirical Constant Threshold (EC-Threshold) h^E is determined for the purpose of controlling the false alarm rate over the sequence of n observations $(X_1, \dots, X_n)_{n \geq 1}$. Knowing that α is the instantaneous false alarm rate tolerated at any given time, we make the approximation that the false alarm rate on the trajectory of n observations is given by $n \cdot \alpha$. In other words, the alarm rate over the whole trajectory of length n is $n \cdot \alpha$. Let us notice that unlike the WC-Threshold h^W , the EC-Threshold depends on the detection objective (δ, q) . The steps to build the EC-Threshold $h^E(\alpha, \delta, q)$ are summarized in Algorithm 1 and Figure 1.

Algorithm 1 Simulation-based procedure to build the EC-Threshold $h^E(\alpha, \delta, q)$

Input: pre-change regime parameters μ_0 and σ_0 , detection objectives δ and q , tolerated false alarm rate α , number of simulated series B , number of observations n .

Initialization: $W_0 = 0$.

for $j = 1$ to B **do**

for $i = 1$ to n **do**

 Generate x_t^j : the t^{th} observation of the j^{th} simulated series, under pre-change regime $\mathbb{P}_0(x_t^j | x_{t-1}^j, \dots, x_1^j)$;

 Compute the score $s_t^j(\delta, q)$;

 Compute $w_t^j(\delta, q) = \max(0, w_{t-1}^j(\delta, q) + s_t^j(\delta, q))$: the corresponding statistic;

end for

 Compute $m^j(\delta, q) = \max_{1 \leq t \leq n} w_t^j(\delta, q)$: the maximum of statistic;

end for

Set the EC-Threshold as the empirical quantile of order $(1 - n\alpha)$ of the B computed maximums:

$$h^E(\alpha, \delta, q) = \mathbf{q}_{(1-n\alpha)} [\{m^j(\delta, q)\}_{1 \leq j \leq B}]$$

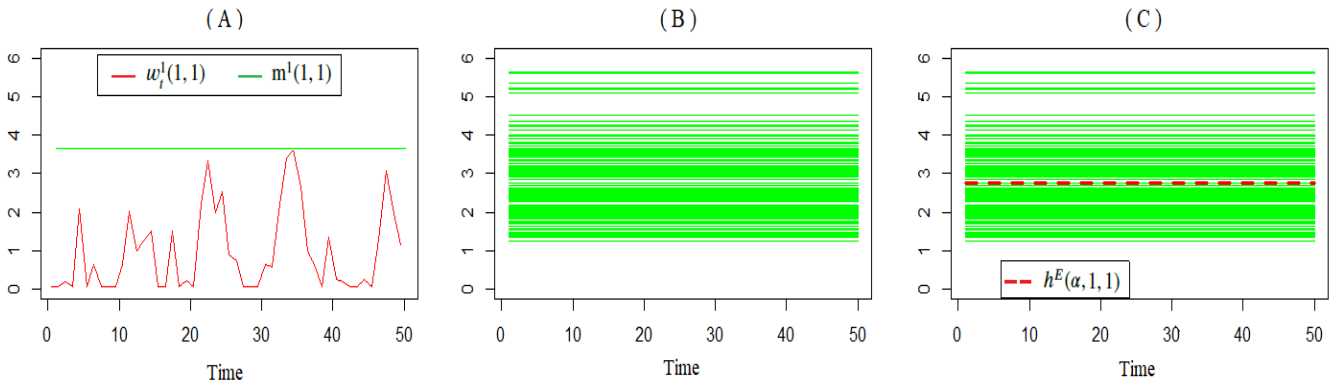


FIGURE 1 Principles for the construction of the EC-Threshold $h^E(\alpha, \delta, q)$ for $\alpha = 0.01$, $\delta = 1$, $q = 1$ and $n = 50$. (A): trajectory of the statistic $w_t^1(1, 1)$ and its maximum $m^1(1, 1)$. (B): computing of the B maximums $\{m^j(1, 1)\}_{1 \leq j \leq B}$. (C): $h^E(0.01, 1, 1)$ is the empirical quantile of order $n\alpha\%$.

Remarks

- The maximum $m^j(\delta, q)$ is called the local score of the series $\{s_t^j(\delta, q)\}_{1 \leq t \leq n}$ and was first introduced by Karlin & Altschul³⁰ in 1990 to study molecular sequences; its distribution was studied by Mercier & Daudin in 2001³¹. Note also that $h^E(\alpha, \delta, q)$ depends on n . We are conscious that our choice of the quantile of order $(1 - n\alpha)$ is a naive choice; further investigations would be interesting.
- It should be noted that the empirical quantile $(1 - n\alpha) \in [0, 1]$, then if n is large, α should be very small to calculate the quantile. For example, for $n = 50$, the tolerated α to be fixed, should not exceed 0.02 ($\alpha \leq 0.02$).

3.3 | Conditional Empirical Instantaneous Threshold

The first ones to use an instantaneous detection threshold are Margavio et al.²⁶. They propose a time-dependent threshold $h_t(\alpha_t)$ to control the conditional false alarm rate α_t at each time t , such that:

$$\begin{cases} \mathbb{P}_0 [W_1(\delta, q) \geq h_1(\alpha_1)] = \alpha_1 \\ \mathbb{P}_0 [W_2(\delta, q) \geq h_2(\alpha_2) | W_1(\delta, q) < h_1(\alpha_1)] = \alpha_2 \\ \vdots \\ \mathbb{P}_0 [W_t(\delta, q) \geq h_t(\alpha_t) | W_i(\delta, q) < h_i(\alpha_i), \forall i < t] = \alpha_t, \quad t \geq 1 \end{cases} \quad (12)$$

In this case, the conditional instantaneous threshold is determined iteratively. At a given time t , to determine $h_t(\alpha_t)$, we should know the tolerated false alarm rates until time t i.e. $(\alpha_1, \alpha_2, \dots, \alpha_t)$ and also the thresholds until time $t - 1$. To simplify, Margavio et al. put α_t constant: $\alpha_t = \alpha, t \geq 1$. Therefore, the probability that a false alarm does not occur at time n under pre-change regime is calculated by:

$$\begin{aligned} \mathbb{P}_0 [T > n] &= \mathbb{P}_0 [W_t(\delta, q) < h_t(\alpha), t = 1, \dots, n] \\ &= (1 - \alpha)^n. \end{aligned}$$

We can see that in this case, T is the first time an alarm is triggered, and it is a random variable that follows a geometric distribution in the discrete case (that would be an exponential distribution in the continuous case). In both cases, the mean time between false alarm is $MTBFA = 1/\alpha$.

By considering (12), we can see that this threshold depends on the objectives of detection δ and q , and it was not the case with the Wald's threshold. To build the conditional instantaneous threshold, we use an empirical method based on simulations such as the one first proposed by Shen et al.²⁹ for a Poisson CUSUM chart, and recently by Huang et al.²⁸ for the generalized CUSUM chart called GLR-CUSUM. The method consists in performing simulations of the statistic under the pre-change regime and in constructing the threshold by the empirical quantile of the statistic. The steps to build this Conditional Empirical Instantaneous Threshold (CEI-Threshold) $h_t^C(\alpha, \delta, q)$ are summarized in Algorithm 2.

Algorithm 2 Simulation-based procedure to build the CEI-Threshold $h_t^C(\alpha, \delta, q)$

Input: pre-change regime parameters μ_0 and σ_0 , detection objectives δ and q , tolerated α , number of simulated series B .

Initialization: $W_0 = 0, t = 1, J_1 = \{1, 2, \dots, B\}$.

repeat

for $j \in J_t$ **do**

 Generate x_t^j : the j^{th} observation at time t , under pre-change regime $\mathbb{P}_0(x_t^j | x_{t-1}^j, \dots, x_1^j)$;

 Compute the score $s_t^j(\delta, q)$;

 Compute $w_t^j(\delta, q) = \max(0, w_{t-1}^j(\delta, q) + s_t^j(\delta, q))$, the corresponding statistic;

end for

 Set the CEI-Threshold at time t as the empirical quantile of order of $(1 - \alpha)$: $h_t^C(\alpha, \delta, q) = \mathbf{q}_{(1-\alpha)}[\{w_t^j(\delta, q)\}_{j \in J_t}]$.

 Let $J_{t+1} = \{j \in J_t; w_t^j(\delta, q) < h_t^C(\alpha, \delta, q)\}$;

 Let $t = t + 1$

until Stop test;

Remarks

- This algorithm differs from ones proposed by Shen et al.²⁹ or Huang et al.²⁸. Ours is more general because it allows the variables $\{x_t^j\}_{t \geq 1}$ to be non independent. For instance, the law of x_t^j could depend on the past values $\{x_s^j\}_{1 \leq s \leq t-1}$ like the Model 2 presented in section 4.1.3.
- Let us note that the size of the set J_t decreases with t because $|J_{t+1}| = \lfloor |J_t|(1 - \alpha) \rfloor \approx (1 - \alpha)^n B$. This is the reason why in the simulations of the Margavio's threshold, B has to be large.
- It is built recursively.

3.4 | New empirical thresholds

In the following, we propose two new empirical thresholds that also depend on the detection objective, built by a simulation-based procedure as in the previous section. We present below the different proposed thresholds: the empirical instantaneous and a data-driven version of it: the dynamic empirical instantaneous. For each of them, we detail the build steps.

3.4.1 | Empirical Instantaneous Threshold

We propose an Empirical Instantaneous Threshold (EI-Threshold) $h_t^I(\alpha, \delta, q)$ which depends on time t . The idea is now to control the probability of false alarm at each time t , by looking for thresholds $h_t^I(\alpha, \delta, q)$ that verify:

$$\mathbb{P}_0 [W_t(\delta, q) \geq h_t^I(\alpha, \delta, q)] = \alpha, \quad t \geq 1. \quad (13)$$

Unlike the CEI-Threshold, the EI-Threshold can be calculated non-recursively or recursively. The steps of the non-recursive building of the EI-Threshold are summarized in Algorithm 3 and Figure 2. The recursive version is detailed in Algorithm 4.

Algorithm 3 Simulation-based procedure to build non-recursively the EI- Threshold $h_t^I(\alpha, \delta, q)$

Input: pre-change regime parameters μ_0 and σ_0 , detection objectives δ and q , tolerated false alarm rate α , number of simulated series B , number of observations n .

Initialization: $W_0 = 0$.

for $j = 1$ to B **do**

for $i = 1$ to n **do**

 Generate x_t^j : the t^{th} observation of the j^{th} simulated series, under pre-change regime $\mathbb{P}_0(x_t^j | x_{t-1}^j, \dots, x_1^j)$;

 Compute the score $s_t^j(\delta, q)$;

 Compute $w_t^j(\delta, q) = \max(0, w_{t-1}^j(\delta, q) + s_t^j(\delta, q))$, the corresponding statistic;

end for

end for

Set the EI- Threshold at time t by taking the empirical quantile of order $(1 - \alpha)$ of the B computed statistics values:

$$h_t^I(\alpha, \delta, q) = \mathbf{q}_{(1-\alpha)} [\{w_t^j(\delta, q)\}_{1 \leq j \leq B}], \quad t = 1, 2, \dots, n$$

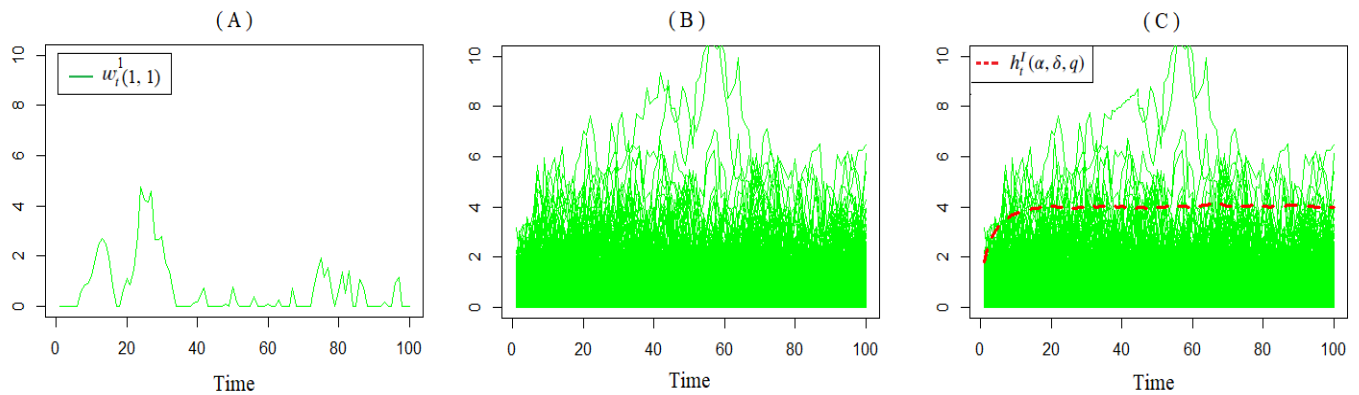


FIGURE 2 Principles for the construction of the EI- Threshold $h_t^I(\alpha, \delta, q)$ until time $n = 100$, for $\alpha = 0.01$, $\delta = 1$, $q = 1$. (A): Trajectory of the statistics $w_t^1(1, 1)$. (B): computing of the $B = 100000$ statistics series $\{w_t^j(1, 1)\}_{1 \leq j \leq B}$. (C): $h_t^I(0.01, 1, 1)$ is the empirical quantile of order of $(1 - \alpha)\%$ at each time t .

Algorithm 4 Simulation-based procedure to build recursively the EI-Threshold $h_t^I(\alpha, \delta, q)$

Input: pre-change regime parameters μ_0 and σ_0 , detection objectives δ and q , tolerated α , number of simulated series B .

Initialization: $W_0 = 0, t = 1$.

repeat

for $j = 1$ to B **do**

 Generate x_t^j : the j^{th} observation at time t , under pre-change regime $\mathbb{P}_0(x_t^j | x_{t-1}^j, \dots, x_1^j)$;

 Compute the score $s_t^j(\delta, q)$;

 Compute $w_t^j(\delta, q) = \max(0, w_{t-1}^j(\delta, q) + s_t^j(\delta, q))$, the corresponding statistic;

end for

 Set the EI- Threshold at time t as the empirical quantile of order of $(1 - \alpha)$: $h_t^I(\alpha, \delta, q) = \mathbf{q}_{(1-\alpha)}[\{w_t^j(\delta, q)\}_{1 \leq j \leq B}]$.

 Let $t = t + 1$

until Stop test;

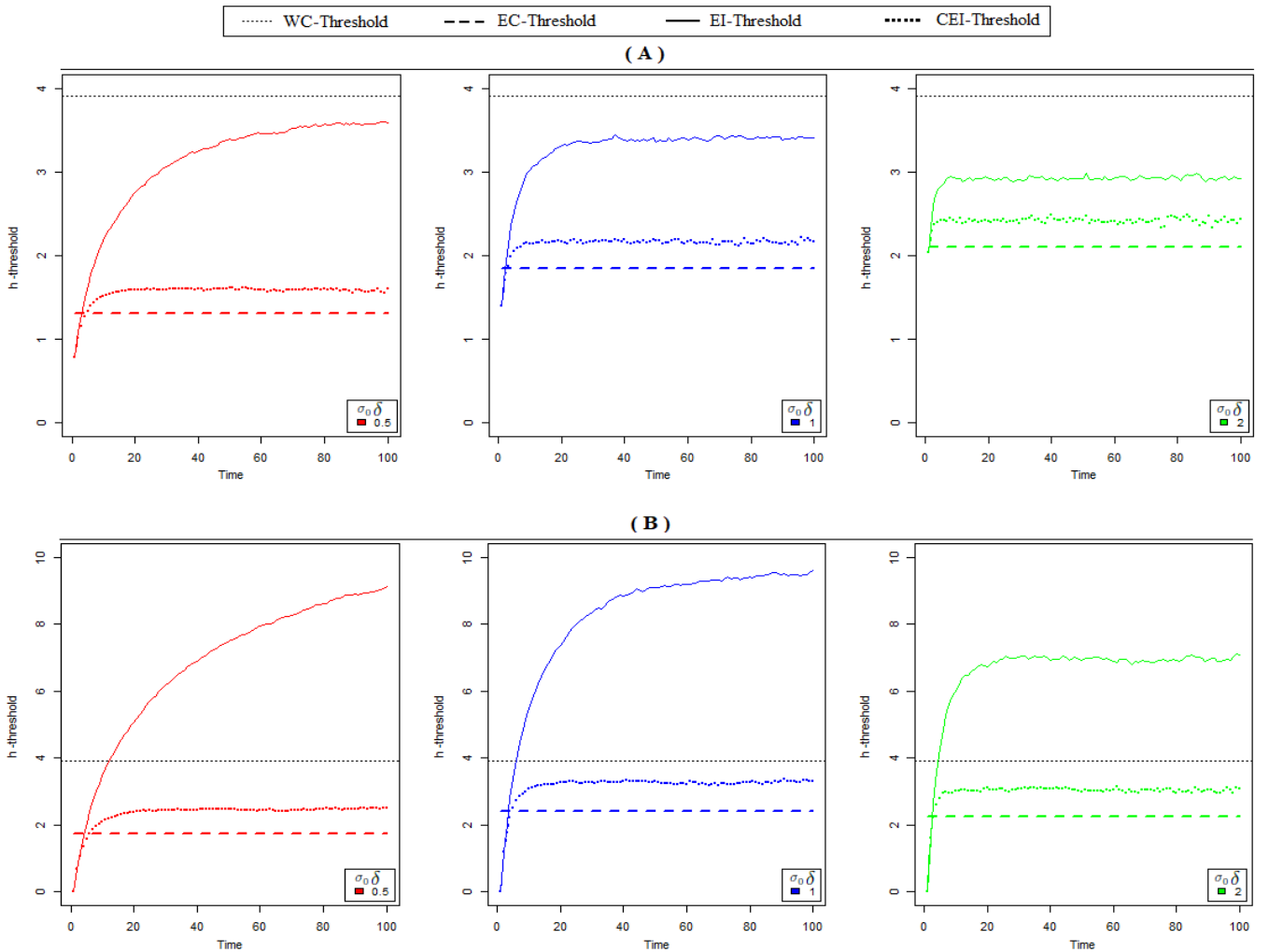


FIGURE 3 Comparison of different thresholds for $\alpha = 0.02$, $(\sigma_0 \delta) \in \{0.5, 1, 2\}$, $q = 1$, with $\sigma_0^2 = \frac{4}{3}$. In the part (A) (resp. (B)), the empirical thresholds are built from i.i.d Gaussian data $\mathcal{N}(0, \sigma_0^2)$ (resp. from an autoregressive model $AR(1)$).

In Figure 3, we show the different thresholds: WC-Threshold h^W (Wald), CEI-Threshold h^C (Margavio), EC-Threshold h^E , EI-Threshold h^I (instantaneous). Thresholds are determined for a tolerated false alarm rate $\alpha = 0.02$, and for three examples of

detection objectives on the mean $(\sigma_0 \cdot \delta) \in \{0.5, 1, 2\}$, $q = 1$, with $\sigma_0^2 = \frac{4}{3}$. In the part (A) - Figure 3, the empirical thresholds are built from independent data of gaussian distribution with $\mu_0 = 0$ and $\sigma_0^2 = \frac{4}{3}$. In the part (B) - Figure 3, the empirical thresholds are built from autocorrelated data of autoregressive processes of order 1 AR(1). We have chosen the parameters in order to have the same mean and variance values in both models.

Firstly, we have to note that the WC-Threshold does not depend on the nature of the data, nor does it depend on time and that it remains the same whatever the expected level of change δ ($h^W(\alpha) = 3.91$). Secondly, we can clearly see that the empirical thresholds constructed for the autoregressive model (B) are considerably higher than those constructed for the independent Gaussian model (A), whatever δ and the threshold are, in particular the EI-Threshold which is the only one to exceed the WC-Threshold.

Let us now focus on part (A) - Figure 3. We can see that the difference between the thresholds is not negligible. WC-Threshold is above the empirical thresholds in this case, so it will detect less quickly, but with fewer false alarms. The EC-Threshold depends on the level of δ ($h^E(\alpha, 0.5, 1) = 1.30$, $h^E(\alpha, 1, 1) = 1.85$, $h^E(\alpha, 2, 1) = 2.10$). It is higher when δ is large. The instantaneous thresholds EI-T and CEI-T, depend strongly on the time at the beginning: we observe a progressive growth until a certain moment, then the thresholds stabilize at a given level. The instant and level of stationarity are specific for each type of threshold and for each detection objective δ . The EI-Threshold is always higher than the CEI-Threshold, and unlike the latter, it is lower when δ is large.

3.4.2 | Dynamical Empirical Instantaneous Threshold

We propose here to make the EI-Threshold h^I , dynamic, that means to adapt it to the behavior of the statistic. The principle is to reset the detection procedure by moving the threshold every time the statistics returns to zero (its initial value). Therefore, the Dynamic Empirical Instantaneous Threshold (DEI-Threshold) denoted h^D is based on the h^I threshold but also depends on the data, in other words, it is "data-driven". This idea is inspired from the principle of Page statistic⁴, which consists in restarting Wald's SPRT algorithm³ whenever the statistic is below the lower threshold (which is suggested to be zero in CUSUM statistic).

Let us suppose that a series $\{X_t\}_{t \geq 1}$ is sequentially observed until time t . The corresponding statistic $W_t(\delta, q)$ is calculated recursively for each new observation. We define by $N_t = \sum_{i=1}^t \mathbf{1}_{\{W_i(\delta, q)=0\}}$ the number of times that the W-statistic has returned to zero at time t . The last time the W statistic is returned to zero can be calculated by the renewal process at each time t by $Z_{N_t} = \inf\{i > Z_{N_t-1}; W_i(\delta, q) = 0\}$. The DEI-Threshold h_t^D is therefore given at each instant t by $h_{t-Z_{N_t}}^I(\alpha, \delta, q)$, and the corresponding stopping time is defined by $T_{h^D} = \min\{t \geq 1 : W_t(\delta, q) \geq h_{t-Z_{N_t}}^I(\delta, q)\}$ with $Z_0 = 0$. The detection procedure by using the DEI-Threshold is summarized in Algorithm 5 and Figure 3.

Algorithm 5 Pseudocode for the detection procedure using the DEI-Threshold $h_t^D(\alpha, \delta, q)$

Input: pre-change regime parameters μ_0 and σ_0 , detection objectives δ and q , tolerated false alarm rate α , EI-Thresholds $h_t^I(\alpha, \delta, q)$.

Initialization: $Z = 0$, $T = \infty$, $w_0(\delta, q) = 0$, $h_1^D(\alpha, \delta, q) = h_1^I(\alpha, \delta, q)$, $t = 1$.

while there is a new observation x_t **do**

Compute $s_t(\delta, q; x_t)$;

Compute $w_t(\delta, q) = \max(0, w_{t-1}(\delta, q) + s_t(\delta, q))$;

if $w_t(\delta, q) = 0$ **then**

$Z = t$;

end if

if $w_t(\delta, q) \geq h_t^D(\alpha, \delta, q)$ **then**

Alarm to signal a detection at time $T = t$;

Go to Initialization step

else

$t = t+1$, and set $h_t^D(\alpha, \delta, q) = h_{t-Z}^I(\alpha, \delta, q)$;

end if

end while

The Figure 4 gives an example from the trajectory of a statistic w_t to show how the EI-Threshold h^I is used to build the data driven dynamical threshold h^D . It has to be noted that before an alarm, the threshold is very frequently reset to the beginning and that in this example, only the eight first points of the EI-Threshold are used in the procedure.

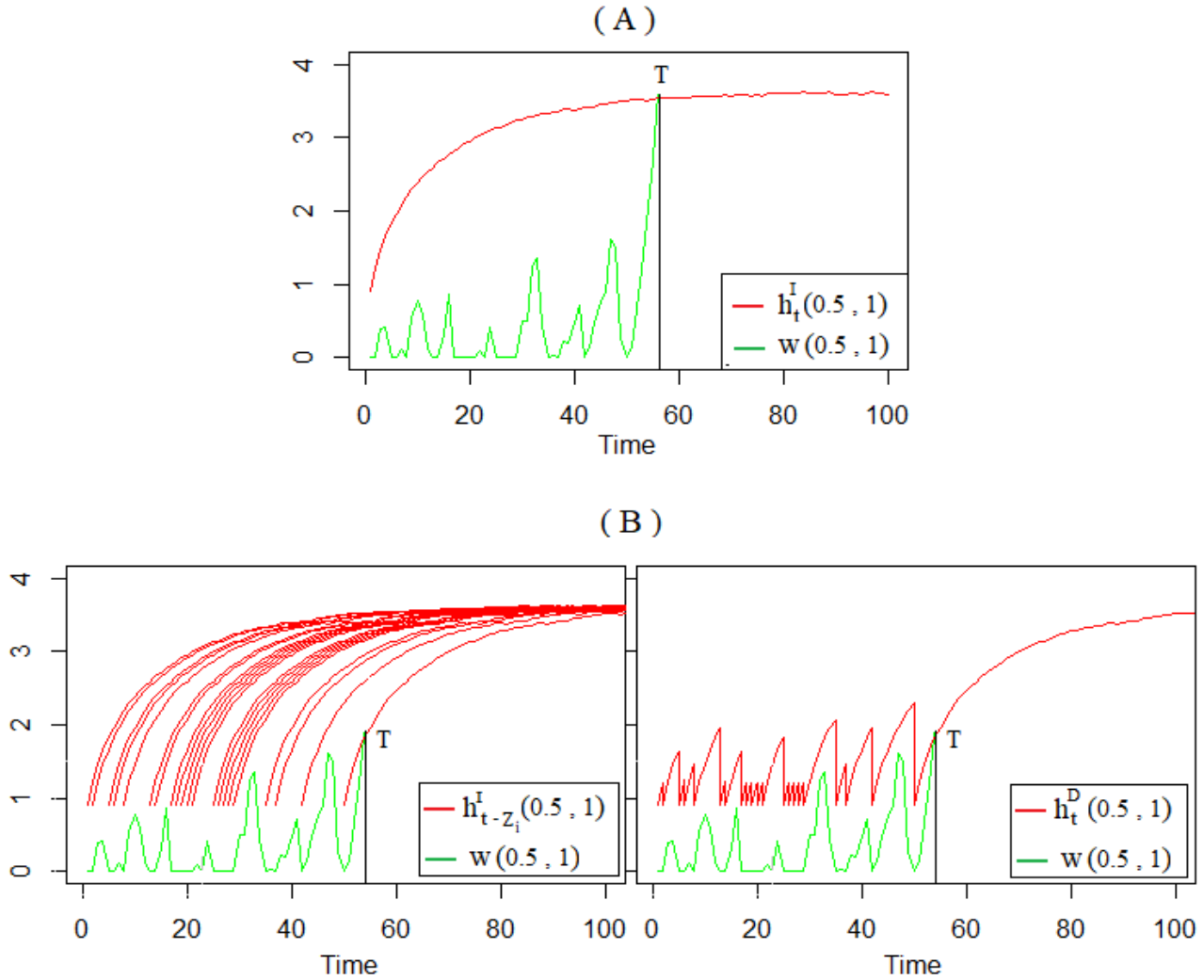


FIGURE 4 Example of calculation of the statistic $w_t(0.5, 1)$ on a simulated series under post-change regime \mathbb{P}_1 , with a change-point at time $\nu = 50$ for a difference level of 0.5 on the mean. (A): detection using the EI-Threshold $h_t^I(0.02, 0.5, 1)$. (B): detection using the DEI-Threshold $h_t^D(0.02, 0.5, 1)$ (data-driven configuration).

4 | THRESHOLDS EVALUATION PROCEDURE

We propose to evaluate the different detection thresholds presented in the previous section on simulated data. To do this, we generate series under pre-change regime \mathbb{P}_0 and post-change regime \mathbb{P}_1 , and estimate for each simulation case the corresponding detection parameters, which are essentially the mean time between false alarm $MTBFA$, the false alarm rate $\alpha = 1/MTBFA$, and the average detection delay ADD .

In what follows, we present the simulation model for the data and the detection objective. Then we will present the procedure used to estimate the detection parameters.

4.1 | Simulation design

In this section, we detail the design of our simulations for the evaluation of the thresholds' performances. Apart Wald threshold, thresholds are built by simulations that depends on an assumed model for the data and on the detection objectives. The assumed model can be different from the one of the data used to evaluated the performance.

4.1.1 | Parameters for the simulation of the data-driven thresholds

For the simulation of the thresholds, we need two kind of parameters: the ones for the simulated data and the ones for the computation of the statistics S_t .

- **Data for the thresholds.** For the building of the thresholds, we have to simulate data under H_0 , we need to know the distribution \mathbb{P}_0 and we denote μ_0, σ_0^2 their expectation and variance. In practice, \mathbb{P}_0 could be unknown and determined by expert opinion or by estimation and could therefore be different from the distribution of the evaluated data \mathbb{P}_0^R (defined in the next section 4.1.2).
- **Parameters for the calculation of the statistic, detection objective.** For the computation of the detection statistic S_t by Equation (4), in addition to μ_0, σ_0^2 (useful to compute $Y_t = \frac{X_t - \mu_0}{\sigma_0}$), we need to know the detection objective parameters $\delta = \frac{\mu_1 - \mu_0}{\sigma_0}$ and $q = \frac{\sigma_1^2}{\sigma_0^2}$. Therefore, we fix δ and q to specify the type and level of change that we would like to detect; they can of course be different from δ^R and q^R (defined in the next section 4.1.2).
- **Length and number of repetitions.** As seen in the different algorithms above, we need also the length n of the series x_t and the number B of repetitions.

4.1.2 | Parameters for the performance's estimation of the thresholds

- **Data for the evaluation.** To perform our simulations of x_t needed for the evaluation of the thresholds, we need to fix the following parameters:
 - Choosing the real pre-change distribution \mathbb{P}_0^R , particularly the choice of the true parameters of the data under pre-change regime μ_0^R and σ_0^R ,
 - Choosing the real post-change distribution \mathbb{P}_1^R , particularly the choice of the true parameters of the data under post-change regime μ_1^R and σ_1^R , that give the real level of change on the mean δ^R and/or on the variance q^R ,
 - Choosing the number n of observations simulated for each trajectory,
 - Choosing the real time of the change-point $1 < v \leq n$ under \mathbb{P}_1^R .
- **Parameters for the calculation of the statistic, detection objective.** For the computation of detection statistic, we take the same parameters as for the thresholds: μ_0, σ_0^2 (useful to compute $Y_t = \frac{X_t - \mu_0}{\sigma_0}$), and detection objective $\delta = \frac{\mu_1 - \mu_0}{\sigma_0}$ and $q = \frac{\sigma_1^2}{\sigma_0^2}$. They can of course be different from δ^R and q^R .
- **Number of repetitions.** For the empirical estimation of α and ADD , we need also to fix B_0 the number of repetitions under H_0 to estimate the false alarm rate and B_1 the number of repetitions under H_1 to evaluate the ADD .

4.1.3 | Models for data

We propose two different models for the data.

Model 1: i.i.d gaussian data

To build the threshold with this model, we used

- Threshold building: $X_t \sim \mathcal{N}(\mu_0; \sigma_0^2)$, $t = 1, \dots, n$.

For the evaluation of the threshold

- Pre-change regime: $X_t \sim \mathcal{N}(\mu_0^R; \sigma_0^{2,R})$, $t = 1, \dots, n$.

- Post-change regime: $X_t \sim \mathcal{N}(\mu_1^R; \sigma_1^{2,R})$, $t = v, \dots, n$.

Model 2 : autocorrelated data

We simulate an autoregressive process of order 1.

To build the threshold with this model, we used

- Threshold building: $X_t = \frac{1}{2}X_{t-1} + \varepsilon_t$, $\varepsilon_t \sim \mathcal{N}(0, 1)$, $t = 1, \dots, n$.
We have to note that $\mathbb{E}_0[X_t] = 0 = \mu_0$ and $\mathbb{V}_0[X_t] = \frac{1}{1-\varphi^2} = \frac{4}{3} = \sigma_0^2$.

For the evaluation of the threshold

- Pre-change regime: $X_t = \frac{1}{2}X_{t-1} + \varepsilon_t$, $\varepsilon_t \sim \mathcal{N}(0, 1)$, $t = 1, \dots, n$.
- Post-change regime: $X_t = Z_t + \sigma_0^R \delta^R$,
with $Z_{v-1} = X_{v-1}$, $Z_t = \frac{1}{2}Z_{t-1} + \varepsilon_t$, $\varepsilon_t \sim \mathcal{N}(0, 1)$, $t = v, \dots, n$.

We have to note that $\mathbb{E}_{0,R}[X_t] = 0 = \mu_0^R$, and $\sigma_0^{2,R} = \sigma_1^{2,R} = \frac{1}{1-\varphi^2} = \frac{4}{3}$ and $\mathbb{E}_{1,R}[X_t] = \sigma_0^R \delta^R = \mu_1^R$.

Remarks

- In the following, we study three cases : the case where the data used to build the thresholds and the ones to evaluate them come both from model 1, the case where the data used to build the thresholds and the ones to evaluate them come both from model 2 and the case where the data used to build the thresholds come from model 1 whereas the data to evaluate them come from model 2.
- For the mean detection case, the pre-change parameters $\mu_0^R = 0$ and $\sigma_0^R = \frac{4}{3}$ in the model 1 and the parameter $\varphi = \frac{1}{2}$ in the model 2 were chosen so as to have a signal of the same expectation and variance on the pre-change regime.
- In any case, our detection objective is to detect a change on the mean of level $(\sigma_0 \delta) \in \{0.5, 1, 2\}$ or on the variance of level $q \in \{0.15, 0.25, 0.50\}$, and to estimate *MTBFA*, α under H_0 , and *ADD* under H_1 with real change-point on the mean of real level $(\sigma_0^R \delta^R) \in \{1, 2\}$ or on the variance of real level $q^R \in \{0.15, 0.25\}$. The real change-point is always set at the time $v = 50$.
- We have chosen $n = 100$ and $B = B_0 = B_1 = 100000$.

4.2 | Parameter Estimation

Parameters *MTBFA* and *ADD* are estimated from the simulation of B independent trajectories $(x_t)_{1 \leq t \leq n}$ on which we sequentially apply our detection procedure. The detection procedure recursively tests the existence of a change up to the time n . At each test output j , either the procedure did not signal a detection, or it stopped at the first detection. In the first case, the alarm time T_j is censored by n , whereas in the second case, the alarm time T_j is observed.

The first part of this section is devoted to the presentation of the procedure used to estimate the *MTBFA* and α under H_0 , the second part, to the procedure to estimate the *ADD* under H_1 .

MTBFA as defined in Equation (7), is the expectation of the alarm instant T . Therefore it should be easy to estimate it by the empirical mean of the times T_j . But we have seen that the data are right censored at the same instant n . We therefore propose to apply a survival analysis. For this purpose, we introduce the following notations:

Let T_j , $j = 1, \dots, B$, be the alarm time of the j^{th} trajectory. For $n \geq 1$, the sample possibly censored information is given by the couple (Z_j, D_j) , $j = 1, \dots, B$, with:

$$Z_j = \min\{T_j, n\}$$

$$D_j = \begin{cases} 1 & \text{if } T_j \leq n : \text{ alarm reported at the time } T_j, \text{ "no censoring"}; \\ 0 & \text{if } T_j > n : \text{ no alarm reported, "censoring"}. \end{cases}$$

We make the hypothesis that the false alarm rate α is constant, so the latent distribution of discrete alarm times is geometric with parameter α . Hence the latent probability density is given by $f_\alpha(z) = \alpha \cdot (1 - \alpha)^{z-1}$, and the associated survival function by $S_\alpha(z) = (1 - \alpha)^z$. The likelihood of the model associated with independent observations $(z_1, d_1), \dots, (z_B, d_B)$ is written in the following form:

$$\begin{aligned} L_{\alpha,n}(z_j, d_j) &= \prod_{j=1}^B \left(\alpha \cdot (1 - \alpha)^{z_j-1} \right)^{d_j} \cdot ((1 - \alpha)^{z_j})^{1-d_j} \\ &= \left(\frac{\alpha}{1 - \alpha} \right)^{\sum_{j=1}^B d_j} \cdot (1 - \alpha)^{\sum_{j=1}^B z_j}. \end{aligned} \quad (14)$$

The Maximum Likelihood Estimator (MLE) of the parameter α is given by

$$\hat{\alpha} = \frac{\sum_{j=1}^B d_j}{\sum_{j=1}^B z_j}. \quad (15)$$

The estimator of α is therefore the ratio between the number of alarms observed and the total exposure to risk. We can deduce immediately the estimator of *MTBFA* by $\widehat{MTBFA} = 1/\hat{\alpha}$.

Remarks

- We observe that in a non-censored model, the *MTBFA* estimator is given by the empirical estimator:

$$\widehat{MTBFA} = \bar{T} = \frac{1}{B} \sum_{j=1}^B t_j.$$

- In the continuous case with exponential latent distribution of parameter α , the estimator of α is also given by:

$$\hat{\alpha} = \sum_{j=1}^B d_j / \sum_{j=1}^B z_j.$$

We now present the estimator of the *ADD* which is given by the expectation of the alarm times under H_1 . Practically speaking, during simulations, each simulated series of n observations has a single change-point at a given time $v \leq n$. There is two cases, either there is a detection at time T_j with a delay $(T_j - v)$, or there is no detection and the delay is censored by $(n - v)$. In this case, the *ADD* estimator using the parametric survival model with a latent geometric distribution is given as follows:

$$\widehat{ADD} = \frac{\sum_{j=1}^B (z_j - v)}{\sum_{j=1}^B d_j} \quad (16)$$

where $z_j = \min\{t_j, n\}$ and $d_j \in \{0, 1\}$.

5 | RESULTS

In this section, we present the results on thresholds' performance obtained by simulation, using the different detection thresholds: WC-Threshold, EC-Threshold, EI-Threshold, DEI-Threshold and CEI-Threshold, according to various models and scenarii of simulation. In section 5.1, we present the simulation results using Model 1 both for the building of the thresholds and the generation of the data. So we first present the case of detection on the mean when the variance is known or correctly estimated

and when it is poorly estimated. Thereafter, we present the case of detection on the variance with a known mean. In section 5.2, we present the results using Model 2 for both building and evaluation of the thresholds. Finally in section 5.3, we present a scenario of a possible error in the estimation of the data distribution between Model 1 and Model 2.

5.1 | Model 1

The results of this section refer to the Model 1 (Section 4.1.3) of i.i.d. Gaussian variables. We successively study the performance of the thresholds (1) for a change in the mean with a known variance, (2) for a change in the mean with an incorrectly estimated variance and (3) for a change in the variance.

5.1.1 | Detection on the mean

We consider Model 1 of i.i.d. Gaussian variables with known variance $\sigma_0^2 = \sigma_0^{2,R} = 4/3$. We give in Figure 5 the estimation results of \widehat{ADD} versus $\hat{\alpha}$ obtained for each of previously described thresholds and for a detection on the mean only. The results are given according to three values of tolerated false alarm rate $\alpha \in \{0.01, 0.02, 0.03\}$ and for different detection objectives $(\sigma_0, \delta) \in \{0.5, 1, 2\}$. The parts (A) and (B) of Figure (5) show the ADD estimation results obtained under the post-change regime, when the real change in the mean is respectively $\sigma_0 \delta^R = 1$ and $\sigma_0 \delta^R = 2$.

First, the results show that the change-point is detected more quickly when the real level of change δ^R is large, whatever the used threshold, the tolerated α or the detection objective (σ_0, δ) . Concerning the thresholds' performance, we generally observe the same behavior of the results whatever the fixed value of α . Let us focus for instance on the case $\sigma_0 \delta^R = 1$ and α fixed to 0.02 (Figure 5, (A), central chart). Depending on the compliance of the threshold with α , it is observed that:

- The EC-Threshold gives the lowest detection delay, but it always exceeds the tolerated α ;
- The CEI-Threshold follows perfectly the fixed α , and reports the best detection delay between those who respect α , $\widehat{ADD}_{hc} = \{4.36, 4.91, 6.11\}$ respectively when $\sigma_0 \delta = \{0.5, 1, 2\}$;
- Among the thresholds that respect the tolerated α , with $\hat{\alpha} < \alpha$, the DEI-Threshold detects faster: $\widehat{ADD}_{hd} = \{6.28, 6.0, 6.86\}$, with $\hat{\alpha}$ close to the fixed one: $\hat{\alpha}_{hd} = \{0.012, 0.015, 0.016\}$, respectively when $\sigma_0 \delta = \{0.5, 1, 2\}$. The greater detection delay is given by the WC-Threshold: $\widehat{ADD}_{hw} = \{12.27, 9.40, 11.25\}$, logically with the lowest false alarm rate: $\hat{\alpha}_{hw} = \{0.001, 0.002, 0.004\}$.

Let us now consider the detection according to different values of δ . We notice that, except for EC-threshold and CEI-threshold, the detection is faster when we ask to detect the exact level of change that exists on the series, ie $\delta = \delta^R$. As we can see in part (A)- Figure 5 when $\sigma_0 \delta^R = 1$, we detect faster for $\sigma_0 \delta = 1$ than for $\sigma_0 \delta = 0.5$; and in part (B)- Figure 5 when $\sigma_0 \delta^R = 2$, we detect faster for $\sigma_0 \delta = 2$ than for $\sigma_0 \delta = 0.5$ and $\sigma_0 \delta = 1$. This results' behavior is valid for the WC-Threshold, EI-Threshold and DEI-Threshold. For these thresholds, we also observed an unexpected behavior of the estimated false alarm rate $\hat{\alpha}$: the more δ is large, the more we have false alarms. On the contrary for the CEI-Threshold and EC-Threshold, for which the detection is usually faster as long as the wanted level of change is less than the real one, and generally $\hat{\alpha}$ is similar regardless of δ .

From these results, one can deduce different facts, the first one is general and almost obvious, the others concern the sensibility of the parameters according to either the objective detection δ , or the real post-change δ^R .

- When δ and δ^R are fixed, we see that the ADD - estimate is a decreasing function of the α -estimate. The slope is very high for the α 's that are near to 0. This result is intuitive since we have to make a compromise between the false alarm rate and the ADD .
- Under H_0 , the EI-Threshold h^I is not consistent in the sense that the estimated rate $\hat{\alpha}_{h^I}$ is highly dependent on the detection objective δ . On the other hand, this inconsistency is reduced when this threshold is made dynamic with DEI-Threshold h^D .
- Under H_1 and particularly when δ^R is large, the WC- and EI-Thresholds give ADD s that strongly depend on the detection objective δ . With the other thresholds, the estimated parameter are not so fluctuating.

Anyway, we have seen that the CEI-Threshold has the best performance because it achieves the expected tolerated false alarm rate and its performance does not depend on the values of the detection objective δ and the real mean change δ^R . Comparing

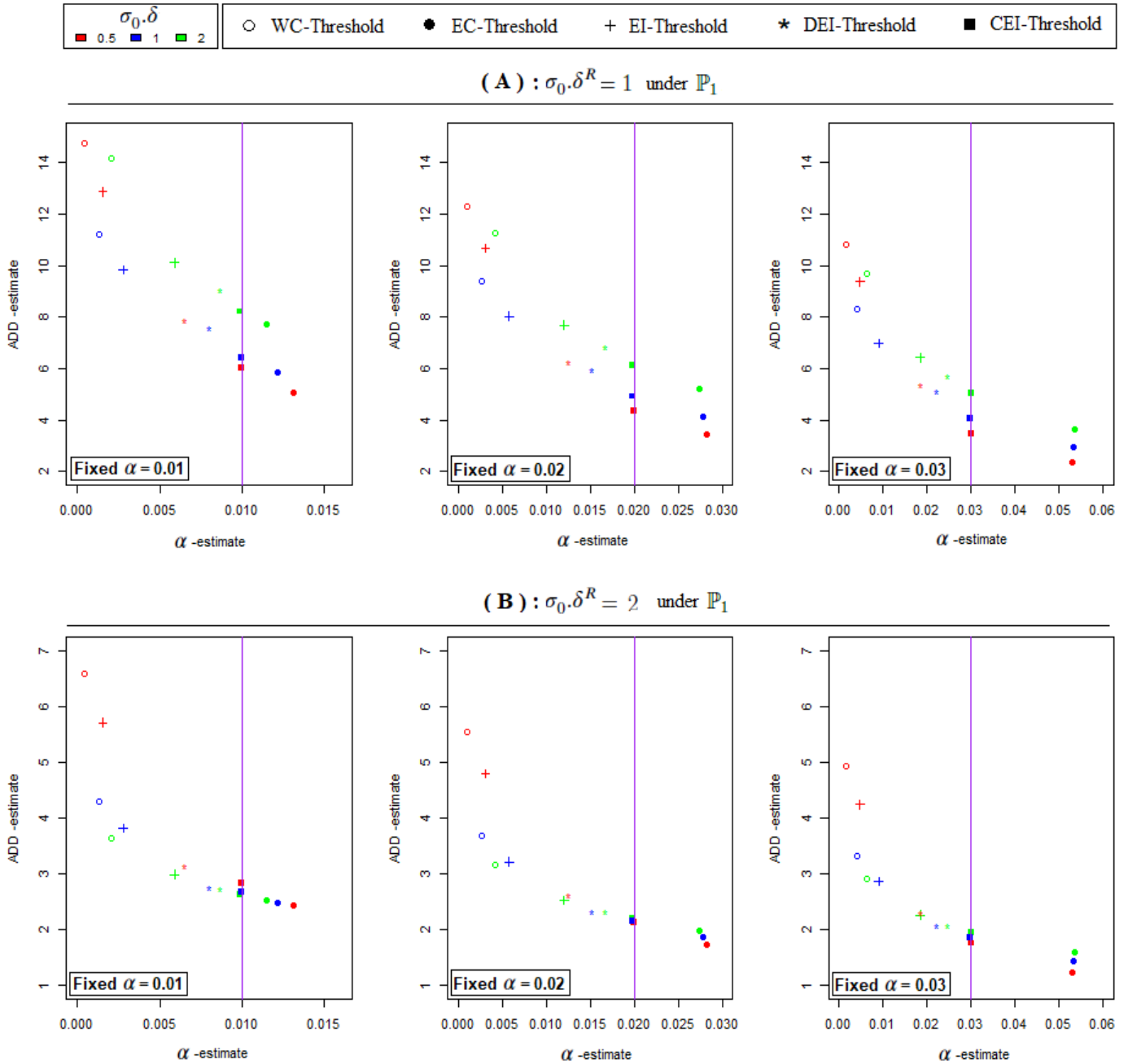


FIGURE 5 Estimation results of \widehat{ADD} vs $\hat{\alpha}$ obtained by the different thresholds for three given values $\alpha \in \{0.01, 0.02, 0.03\}$ for Model 1 (i.i.d. Gaussian) and according to $\sigma_0 \delta \in \{0.5, 1, 2\}$, with $\sigma_0^2 = \sigma_0^{2,R} = \frac{4}{3}$. In the part (A) (resp. (B)), the \widehat{ADD} s are estimated under H_1 with real change from the mean of level of $\sigma_0 \delta^R = 1$ (resp. $\sigma_0 \delta^R = 2$).

to the other thresholds, the DEI-Threshold gives interesting results. It respects indeed the expected false alarm rate and realizes a detection delay relatively close to the CEI-Threshold's.

5.1.2 | Detection on the mean, with a mis-estimated variance

We still consider Model 1 of i.i.d. Gaussian variables for the detection of a change in the mean only. We suppose that parameters $\sigma_0 = \sigma_1$ used for the computation of the statistic $S_t(\delta, q)$, are not the real ones denoted by $\sigma_0^R = \sigma_1^R$. In Figure 6, we compare

the estimation results obtained by the different thresholds for a detection on the mean from three different situations: (1) the real variance of the signals is lower than the estimated one ($\sigma_0^{2,R} = \frac{1}{2} < \sigma_0^2 = \frac{4}{3}$), (2) the real variance of the signals equals the fixed one ($\sigma_0^{2,R} = \sigma_0^2 = \frac{4}{3}$), (3) the real variance of the signals is greater than the estimated one ($\sigma_0^{2,R} = 2 > \sigma_0^2 = \frac{4}{3}$). We have given the results for the tolerated false alarm rate $\alpha = 0.02$, and various detection objective $\sigma_0\delta \in \{0.5, 1, 2\}$. Under H_1 , the real change on the mean is of level $\sigma_0^R\delta^R = 1$.

The results show that a bad estimation of the variance parameter, with these observed levels, does not affect detection too much and does not change the trend of the results for the different thresholds. With an underestimation of the variance, we may get more false alarms and potentially exceed the tolerated rate, as shown for the EI-, CEI- and DEI- thresholds, but with a slightly faster detection for all thresholds. With an overestimation of the variance, we may detect the change-point a little more slowly for every threshold, but in this case we will gain much less false alarms, and all the thresholds respect the tolerated rate.

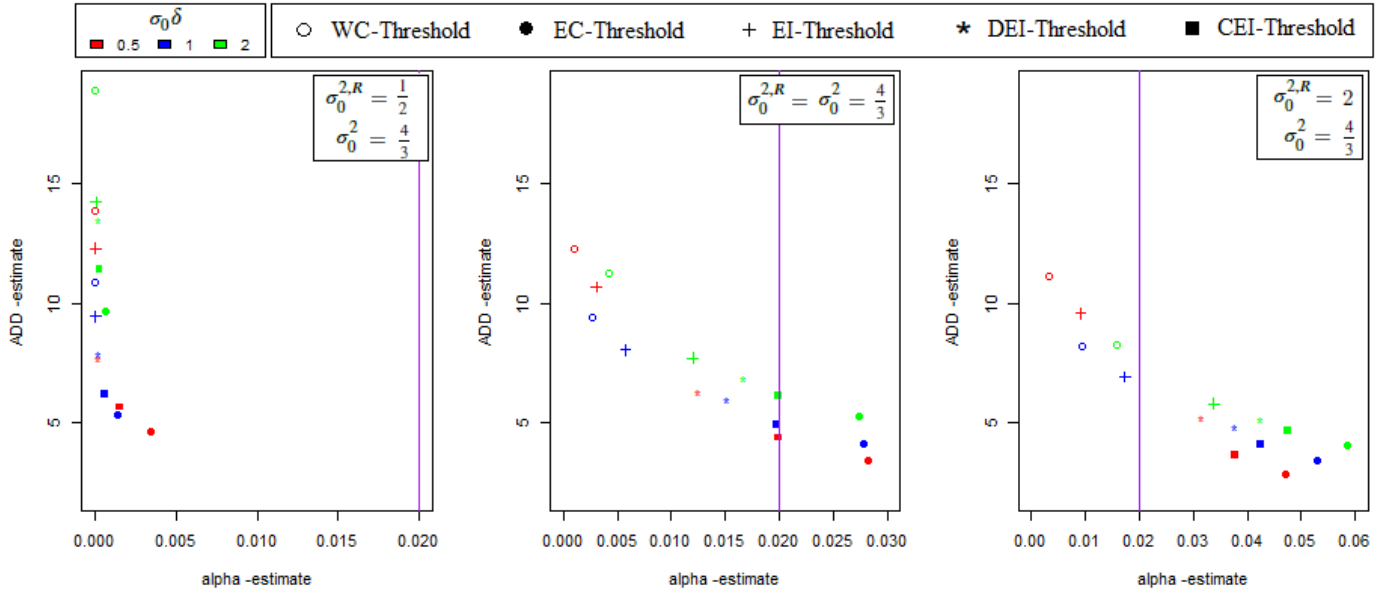


FIGURE 6 Estimation results of \widehat{ADD} vs $\hat{\alpha}$ obtained by the different thresholds for the value $\alpha = 0.02$, for data from Model 1 (i.i.d. Gaussian) and according to $\sigma_0\delta \in \{0.5, 1, 2\}$, with $q = 1$ and $\sigma_0^2 = \frac{4}{3}$. From left to right, we compare the results for three cases: $\sigma_0^{2,R} = \frac{1}{2} < \sigma_0^2$, $\sigma_0^{2,R} = \frac{4}{3} = \sigma_0^2$ and $\sigma_0^{2,R} = 2 > \sigma_0^2$. The \widehat{ADD} s are estimated with real change in mean of level $\sigma_0^R\delta^R = 1$.

5.1.3 | Detection on the variance

In Figure 7, we present the estimation results of \widehat{ADD} versus $\hat{\alpha}$ obtained by each threshold for a change-detection on the variance i.e. with $\mu_0^R = \mu_0 = \mu_1^R = \mu_1 = 0$, according to three values of tolerated false alarm rate $\alpha \in \{0.01, 0.02, 0.03\}$ and for different detection objective $q \in \{0.15, 0.25, 0.50\}$. The parts (A) (resp. (B)) of Figure (7) gives the ADD estimation results obtained under the post-change regime H_1 , when the real change in the variance is $q^R = 0.25$ (resp. $q^R = 0.15$).

For the purpose of comparing the threshold's performance, we globally observe the same behavior than the detection on the mean. Although we can remark some slight differences depending on the level of change on the variance which we ask to detect (here, the smaller is q , the greater is the level of change):

- The change-point is detected faster if the requested level of change is less than or equal to the actual change ($q \geq q^R$), except for EI- and DEI-Thresholds.
- For the EI-Threshold, the detection delay is large when q is large (the level of change requested is smaller than the real one in the signals). In contrast, when using the dynamic configuration of the EI-Threshold, the detection delay does not depend on the requested detection objective q . The \widehat{ADD} results of the DEI-Threshold are consistent.

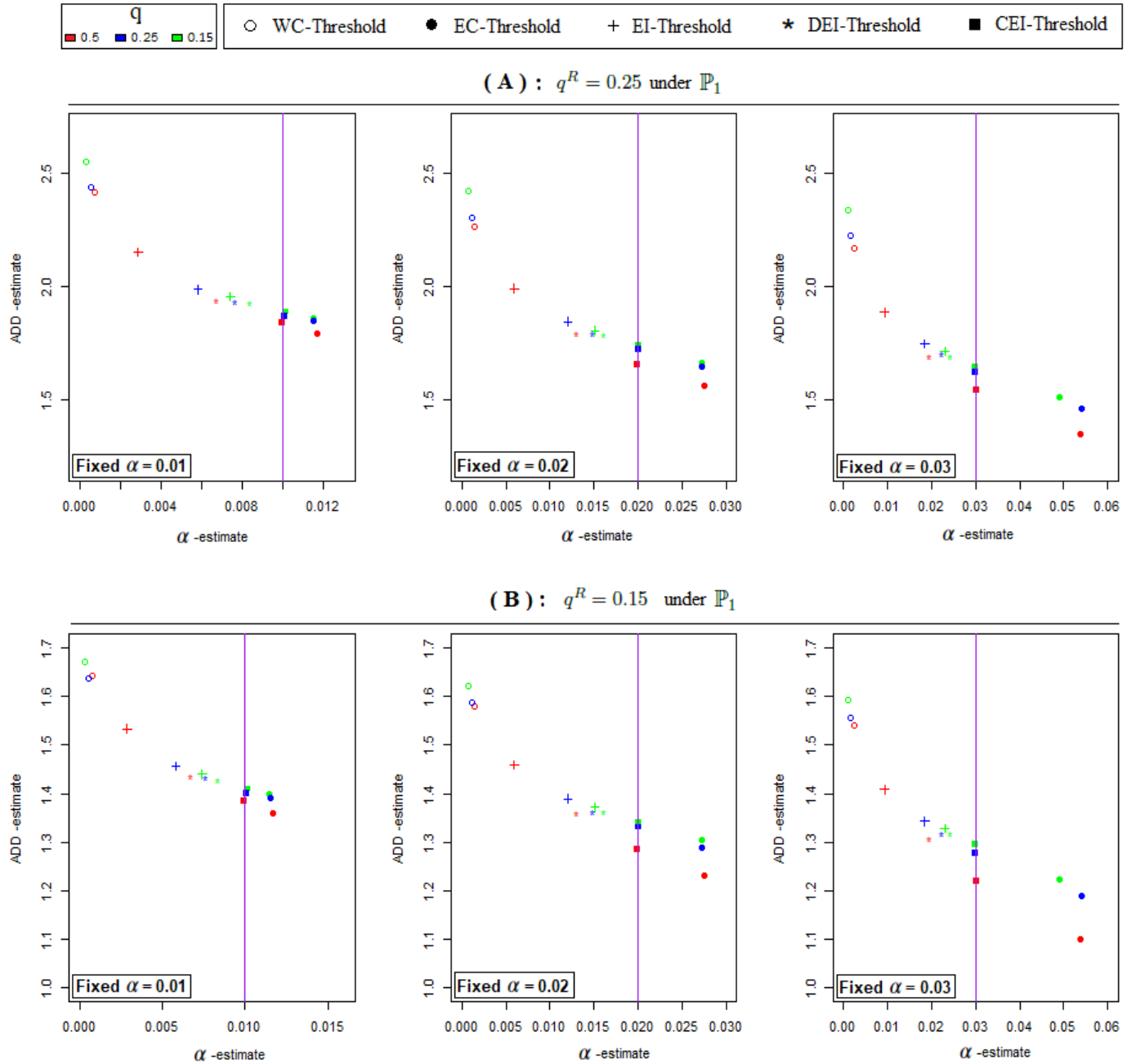


FIGURE 7 Estimation results of \widehat{ADD} vs $\hat{\alpha}$ obtained by the different thresholds for i.i.d. Gaussian data (Model 1), three given values $\alpha \in \{0.01, 0.02, 0.03\}$, and $q \in \{0.15, 0.25, 0.50\}$, with $\delta = 0$ and $\mu_0 = \mu_0^R = 0$. In the part (A) (resp. (B)), the \widehat{ADD} s is estimated with real change in variance of level of $q^R = 0.25$ (resp. $q^R = 0.15$).

5.2 | Model 2

In this section, we choose autocorrelated simulated data. Details on the used autoregressive process are given in section 4.1.3. Here we assume that the data distribution is known or correctly estimated. We also used this data generation process to build the empirical thresholds.

We give in Figure 8 the estimation results of \widehat{ADD} versus \hat{a} obtained by the various thresholds for a detection on the mean i.e. $q = 1$ with $\sigma_0^2 = \sigma_0^{2,R} = \frac{4}{3}$. The results are given according to three values of tolerated false alarm rate $\alpha \in \{0.01, 0.02, 0.03\}$ and for different detection objective $\sigma_0\delta \in \{0.5, 1, 2\}$. In parts (A) (resp. (B)) of Figure (8), the ADD estimation under the post-change regime are obtained when the real change in the mean is respectively $\sigma_0\delta^R = 1$ (resp. $\sigma_0\delta^R = 2$).

For a fixed tolerated false alarm rate α and the real change $\sigma_0\delta^R$, comparing results of model 1 (Figure 8) and of model 2 (Figure 5) show that:

- The detection delay given by every threshold and according to $\sigma_0\delta$, is greater in model 2 than in model 1, especially for the EI-Threshold, with the exception of the WC-Threshold which is slightly faster in model 2, logically because the WC-Threshold is the same in both models. It is not empirical.
- The thresholds' performance trend is generally the same in both models, except that the WC-Threshold gives better results in model 2, relatively close to those of the CEI-Threshold.
- In general, we have the same behavior according to $\sigma_0\delta$, except for the DEI-Threshold which exceeds the tolerated α when $\sigma_0\delta = 2$.

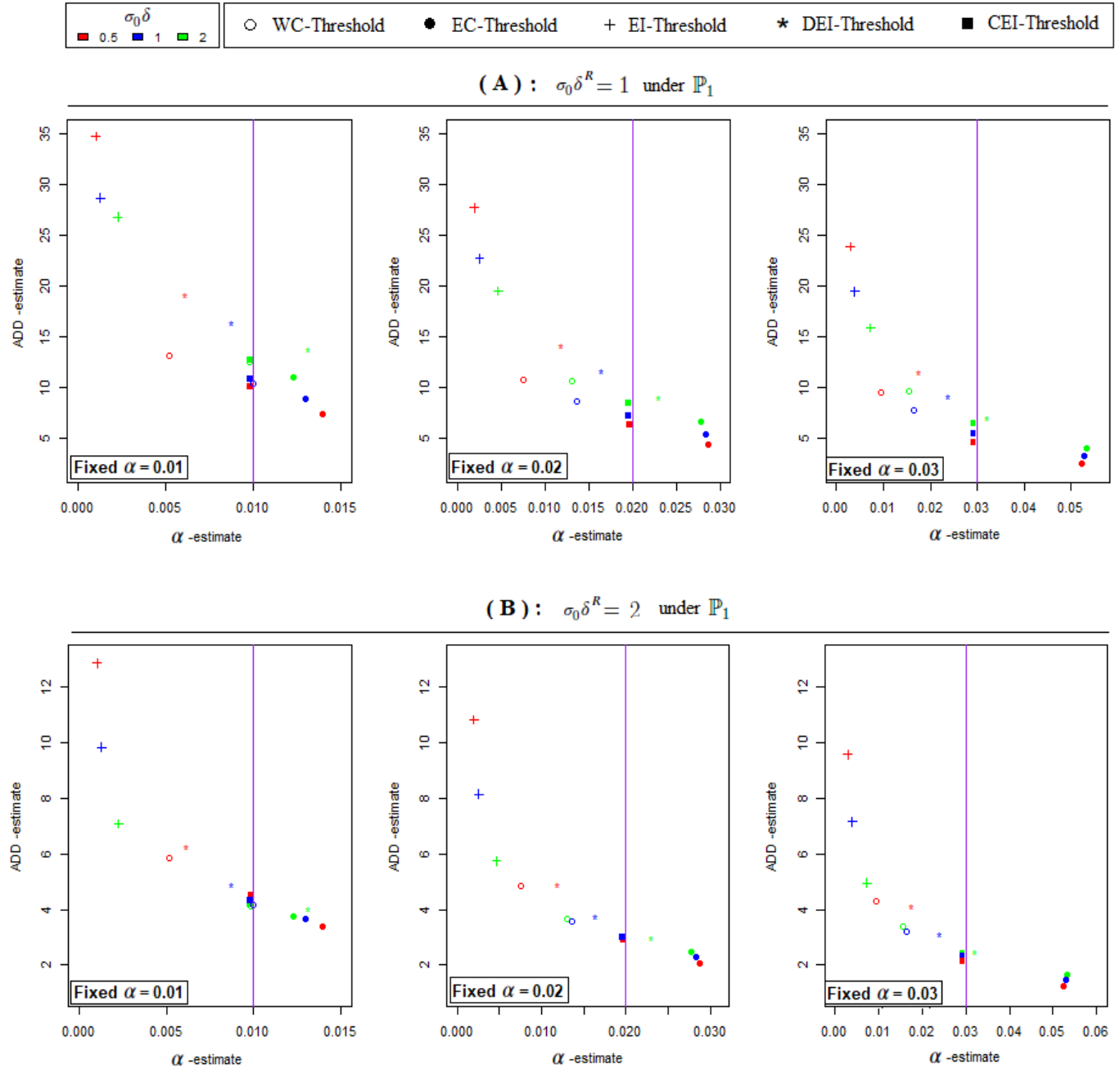


FIGURE 8 Estimation results of \widehat{ADD} vs $\hat{\alpha}$ obtained by the different thresholds for autocorrelated data (Model 2), three given values $\alpha \in \{0.01, 0.02, 0.03\}$, $\sigma_0 \delta \in \{0.5, 1, 2\}$, and $\sigma_0^2 = \frac{4}{3}$. In the part (A) (resp. (B)), the \widehat{ADD} estimated respectively with real change of level of $\sigma_0 \delta^R = 1$ (resp. $\sigma_0 \delta^R = 2$).

5.3 | Distribution estimation error

In this section, we present the threshold's performance obtained in the following situation where the distribution is misestimated: we use for detection the empirical thresholds constructed by model 1, on data simulated from model 2.

In Figure 9, we give the estimation results of \widehat{ADD} versus $\hat{\alpha}$ obtained by each threshold for a detection on the mean i.e. $q = 1$ with $\sigma_0^2 = \sigma_0^{2,R} = \frac{4}{3}$. The results are given according to three values of tolerated false alarm rate $\alpha \in \{0.01, 0.02, 0.03\}$ and for different detection objective $\sigma_0\delta \in \{0.5, 1, 2\}$. The ADD estimation under the post-change regime are obtained when the real change in the mean is $\sigma_0\delta^R = 1$.

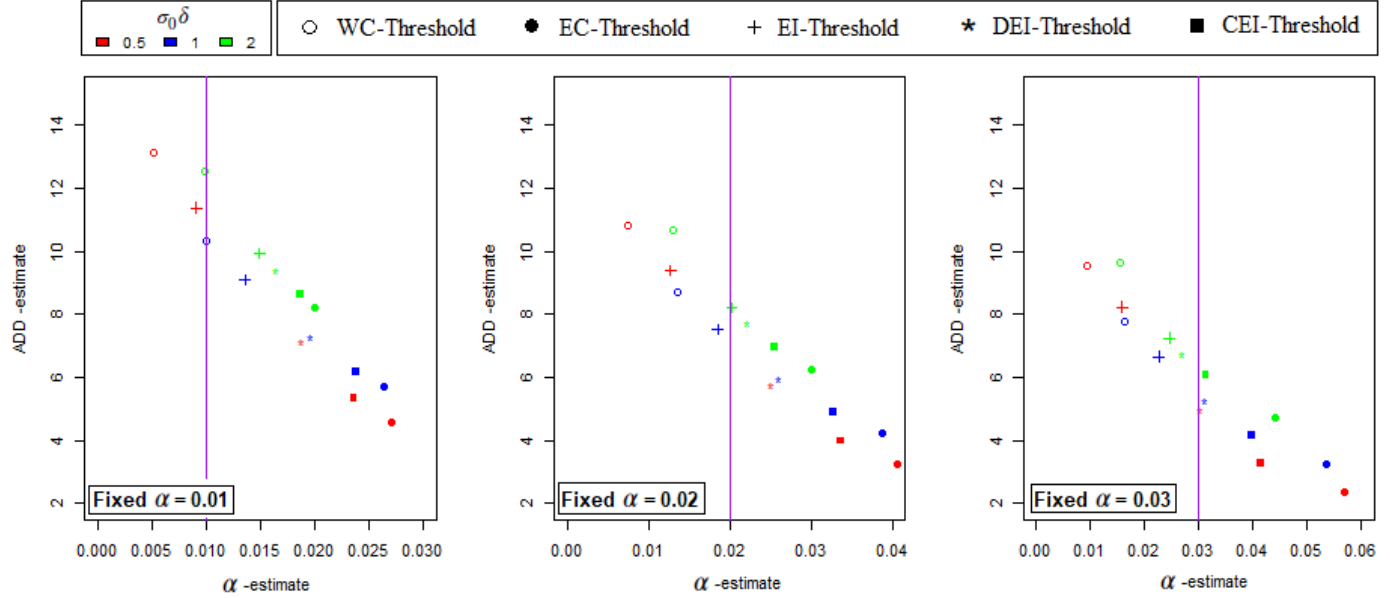


FIGURE 9 Estimation results of \widehat{ADD} vs $\hat{\alpha}$ obtained by the different thresholds built from i.i.d. Gaussian data (Model 1), for autocorrelated data (Model 2), for three given values $\alpha \in \{0.01, 0.02, 0.03\}$, and $\sigma_0\delta \in \{0.5, 1, 2\}$, with $\sigma_0^2 = \frac{4}{3}$. The \widehat{ADD} s are estimated with real change value $\sigma_0\delta^R = 1$.

The Figure 9 shows performance results for a misevaluation of the data distribution, that is, if we assume the independence of the data and build the empirical thresholds under an i.i.d Gaussian distribution, while the data are autocorrelated. In this case we run the risk of having more false alarms than usual and logically a little faster detection, for every threshold. Therefore this may lead to a false alarm rate that exceeds the tolerated one, in particular when α is small. For example, when the distribution used for the threshold is the one of the data, the estimated false alarm rate of the CEI- and DEI-Thresholds respect always α , but here they don't do it any longer: we gain a shorter detection delay than usual. For a fixed detection objective δ , the order of performances of the different thresholds remains the same. We also note, in particular for the EC-, DEI- and CEI-Thresholds, that when $\sigma_0\delta = 2$, we have significantly fewer false alarms than for $\sigma_0\delta = 0.5$ and $\sigma_0\delta = 1$.

6 | CONCLUSION AND PERSPECTIVES

We have dealt with the case of change-point detection in the sequential context where the challenge is to detect a change as quickly as possible while respecting a given false alarm rate. We considered the CUSUM recursive statistic based on a score function.

The aim of this article is to compare the performance of different detection thresholds. Some of them come from the literature (Wald's Constant, Empirical Constant and the Margavio's Conditional Empirical Instantaneous) and the others, built by a simulation-based procedure, are original contributions (Empirical Instantaneous and Dynamic Empirical Instantaneous). The

detection performances are evaluated on simulated data by estimating mainly the conditional false alarm rate α and the average detection delay *ADD* under respectively hypothesis H_0 and alternative hypothesis H_1 involving post-change regime. For a correct estimation of these parameters, we used survival analysis because the observations are censored by the observation limit. The survival model is built under the hypothesis of a constant conditional alarm rate.

The different detection thresholds were assessed under two main simulation models. In the first one, the data are independent from Gaussian distribution. In the second one, the data are autocorrelated from an autoregressive process. Under the first model, we looked at shift detection on the mean and on the variance. For a shift detection on the mean, we also evaluated the sensitivity of the findings in the case where the signal variance parameter is not accurately specified. Finally, we proposed to assess the impact of an error in the choice of data distribution, with thresholds constructed on the first simulation model but applied to data of the second simulation model.

The results showed that :

- The performances of the different thresholds for detection on the mean and on the variance are comparable. We have the same behavior and trend of results.
- The empirical thresholds are more efficient (smaller delay) in the first model where the data are independent than in the second one where they are not. In contrast, the non-empirical WC-Threshold is relatively more efficient and competitive in the second model.
- The thresholds performance is not very sensitive to an incorrect estimation of the variance parameter since the speed of detection is not really affected. However, we risk an increase (respectively a decrease) of false alarms in the case of underestimation (respectively overestimation).
- The performance of the thresholds seems to be robust even if an error is made in the choice of the data distribution.
- The CEI-Threshold and the DEI-Threshold give the best performances on the simulated data. They establish a compromise between the detection delay and the false alarm level. The CEI-Threshold outperforms the DEI-Threshold as it minimizes the average detection delay *ADD*.

If we analyse the results of this study in the perspective of applying the simulation-based CEI- or DEI-Thresholds on real data, we have to keep in mind that to generate simulation-based thresholds, the pre-change distribution has to be determined. This should be done from the data. Note that in real cases the series of observations is probably not a sequence of independent variables. From a practical point of view, the DEI-Threshold is much easier to determine because, as we have seen in Figure 4, it uses only the beginning of the EI-Threshold path. It is therefore very easy to compute and to update with new information.

The perspective of this work is to use the DEI-Threshold in order to detect rupture in real signal, in the context of e-health for patients followed by sensors. The beginning of the DEI-Threshold will be estimated from the data under the normal (pre-change) regime and updated along the process.

ACKNOWLEDGMENTS

We would like to thank the reviewers for their fruitful comments which greatly improved the article.

We thank Sabine MERCIER for the fruitful discussions on the local score and on the link between the simulation-based thresholds EC and DEI with the current excursions of the CUSUM statistic.

References

1. Shewhart WA. *Economic control of quality of manufactured product*. Macmillan And Co Ltd, London . 1931.
2. Neyman J, Pearson ES. IX. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 1933; 231(694-706): 289–337.
3. Wald A. Sequential tests of statistical hypotheses. *The annals of mathematical statistics* 1945; 16(2): 117–186.
4. Page ES. Continuous inspection schemes. *Biometrika* 1954; 41(1/2): 100–115.
5. Lorden G. Procedures for reacting to a change in distribution. *The Annals of Mathematical Statistics* 1971; 42(6): 1897–1908.
6. Moustakides GV. Optimal stopping times for detecting changes in distributions. *The Annals of Statistics* 1986; 14(4): 1379–1387.
7. Shiryaev A. The problem of the most rapid detection of a disturbance in a stationary process. In: . 2. ; 1961.
8. Shiryaev AN. On optimum methods in quickest detection problems. *Theory of Probability & Its Applications* 1963; 8(1): 22–46.
9. Roberts S. A comparison of some control chart procedures. *Technometrics* 1966; 8(3): 411–430.
10. Pollak M. Optimal detection of a change in distribution. *The Annals of Statistics* 1985: 206–227.
11. Polunchenko AS, Tartakovsky AG, others . On optimality of the Shiryaev–Roberts procedure for detecting a change in distribution. *The Annals of Statistics* 2010; 38(6): 3445–3457.
12. Basseville M, Nikiforov IV. *Detection of abrupt changes: theory and application*. 104. Prentice Hall Englewood Cliffs . 1993.
13. Poor HV, Hadjiliadis O. *Quickest detection*. Cambridge University Press . 2008.
14. Tartakovsky A, Nikiforov I, Basseville M. *Sequential analysis: Hypothesis testing and changepoint detection*. Chapman and Hall/CRC . 2014.
15. Lai TL. Sequential changepoint detection in quality control and dynamical systems. *Journal of the Royal Statistical Society: Series B (Methodological)* 1995; 57(4): 613–644.
16. Tartakovsky AG, Rozovskii BL, Blažek RB, Kim H. A novel approach to detection of intrusions in computer networks via adaptive sequential and batch-sequential change-point detection methods. *IEEE Transactions on Signal Processing* 2006; 54(9): 3372–3382.
17. Tartakovsky AG, Rozovskii BL, Blažek RB, Kim H. Detection of intrusions in information systems by sequential change-point methods. *Statistical methodology* 2006; 3(3): 252–293.
18. Tartakovsky AG, Polunchenko AS, Sokolov G. Efficient computer network anomaly detection by changepoint detection methods. *IEEE Journal of Selected Topics in Signal Processing* 2012; 7(1): 4–11.
19. Tartakovsky AG. Rapid detection of attacks in computer networks by quickest changepoint detection methods. In: World Scientific. 2014 (pp. 33–70).
20. Pepelyshev A, Polunchenko AS. Real-time financial surveillance via quickest change-point detection methods. *arXiv preprint arXiv:1509.01570* 2015.
21. Bakhache B, Nikiforov I. Reliable detection of faults in measurement systems. *International Journal of adaptive control and signal processing* 2000; 14(7): 683–700.

22. Egea-Roca D, Seco-Granados G, López-Salcedo JA. Comprehensive overview of quickest detection theory and its application to GNSS threat detection. *Gyroscope and Navigation* 2017; 8(1): 1–14.
23. Yu X, Baron M, Choudhary PK. Change-point detection in binomial thinning processes, with applications in epidemiology. *Sequential Analysis* 2013; 32(3): 350–367.
24. Sonesson C, Bock D. A review and discussion of prospective statistical surveillance in public health. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 2003; 166(1): 5–21.
25. Tsiamyrtzis P, Hawkins DM. Bayesian statistical process control for Phase I count type data. *Applied Stochastic Models in Business and Industry* 2019; 35(3): 766–787.
26. Margavio TM, Conerly MD, Woodall WH, Drake LG. Alarm rates for quality control charts. *Statistics & Probability Letters* 1995; 24(3): 219–224.
27. Hawkins DM, Qiu P, Kang CW. The changepoint model for statistical process control. *Journal of quality technology* 2003; 35(4): 355–366.
28. Huang W, Shu L, Woodall WH, Tsui KL. CUSUM procedures with probability control limits for monitoring processes with variable sample sizes. *IIE Transactions* 2016; 48(8): 759–771.
29. Shen X, Zou C, Jiang W, Tsung F. Monitoring Poisson count data with probability control limits when sample sizes are time varying. *Naval Research Logistics (NRL)* 2013; 60(8): 625–636.
30. Karlin S, Altschul SF. Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proceedings of the National Academy of Sciences* 1990; 87(6): 2264–2268.
31. Mercier S, Daudin JJ. Exact distribution for the local score of one iid random sequence. *Journal of Computational Biology* 2001; 8(4): 373–380.
32. Moustakides GV, Polunchenko AS, Tartakovsky AG. Numerical comparison of CUSUM and Shiryaev–Roberts procedures for detecting changes in distributions. *Communications in Statistics- Theory and Methods* 2009; 38(16-17): 3225–3239.
33. Hadjiliadis O, Zhang H, Poor HV. One shot schemes for decentralized quickest change detection. *IEEE Transactions on Information Theory* 2009; 55(7): 3346–3359.

AUTHOR BIOGRAPHY

Nassim Sahki is Phd Student in Applied Mathematics and Statistics at the Université de Lorraine, in the Inria team BIology Genetics and Statistics (BIGS). He earned his Master's Degree in Mathematics, Computing and Applications at the Université Grenoble Alpes. He is interested in data science and statistics as applied to the health sciences. He is member of European Network for Business and Industrial Statistics (ENBIS).

Anne Gégout-Petit is full professor of Statistics at Université de Lorraine in France. She received her PhD in mathematics at the Université de Provence. Her research area is mathematical statistics, modelling with application to reliability and life science. She is the head of the Inria team BIology Genetics and Statistics (BIGS). She is member of European Network for Business and Industrial Statistics (ENBIS).

Sophie Wantz-Mézières is an assistant professor in the Probability and Statistics team of IECL, Université de Lorraine (France). She received her PhD in Applied Mathematics at the Université de Lorraine, on Stochastic Differential Equations. Her current research interests are Applied Statistics, in particular applications to medical topics. She belongs to the Inria team BIGS. She is a member of the European Network for Business and Industrial Statistics (ENBIS).

